

| GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd. | | | | | | | | | |
|--|--------|---------------|----|--|---|---|------|---------|--|
| DM protein - protein search, using sw model | | | | | | | | | |
| Run on: March 1, 2001, 16:16:31 ; Search time 91.75 seconds (without alignments) | | | | | | | | | |
| Perfect score: 355 US-09-331-631A-1_COPY_186_248 Sequence: 1 KRPQQQREYEDRRCEQE. MNNFQRGSSGRYYEEEQS 63 | | | | | | | | | |
| Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5 | | | | | | | | | |
| Searched: 88757 seqs., 32294092 residues | | | | | | | | | |
| Minimum DB seq length: 0 Maximum DB seq length: 2000000000 | | | | | | | | | |
| Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries | | | | | | | | | |
| database : SwissProt_39; * | | | | | | | | | |
| Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution. | | | | | | | | | |
| 8 SUMMARIES | | | | | | | | | |
| Result No. Score Query Length DB ID Description | | | | | | | | | |
| 1 114 32.1 588 1 VCLB_GOSHI | P03801 | gossypium h | AC | P09801; | 01-MAR-1989 (Rel. 10, Created) | STANDARD; | PRT; | 588 AA. | |
| 2 101 28.5 605 1 VCLA_GOSHI | P03799 | gossypium h | CC | DT 01-MAR-1989 (Rel. 10, Last sequence update) | DT 15-JUL-1999 (Rel. 38, Last annotation update) | | | | |
| 3 84 23.7 966 1 SSN6_YEAST | P14922 | saccharomye | CC | VICILIN C72 PRECURSOR (ALPHA-GLOBULIN B) | Gossypium hirsutum (Upland cotton). | | | | |
| 4 79 22.3 919 1 ANDR_HUMAN | P10275 | homo sapien | CC | CC | Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Malvales; Malvaceae; Gossypium | | | | |
| 5 75 1090 1 NITA_NEUCR | P8349 | neurospora | CC | CC | Malnoiophyta; Eudicotyledons; Core eudicots; Rosidae; eurosid II; Malvales; Malvaceae; Gossypium | | | | |
| 6 74 20.8 1898 1 TRRY_HUMAN | Q07283 | homo sapien | CC | CC | [1] | | | | |
| 7 74 20.8 242 1 CBP_HUMAN | Q92793 | homo sapien | CC | CC | SEQUENCE FROM N.A. | | | | |
| 8 73.5 20.7 1 T2_MOUSE | Q08666 | mus musculu | CC | CC | Chian C.A., Pyle J.B., Iegocki A.B., Dure L. III; | | | | |
| 9 73 20.6 573 1 GLBL_MATEZ | P15590 | zea mays (m | CC | CC | "developmental biochemistry of cotyledon seed embryogenesis and germination. XVIII. cDNA and amino acid sequences of the members of the storage protein families." | | | | |
| 10 73 20.6 911 1 ANDR_PANTR | Q97775 | pan troglod | CC | CC | Plant Mol. Biol. 7:475-489 (1986). | | | | |
| 11 70.5 19.9 1344 1 IF3A_MOUSE | P33116 | mus musculu | CC | CC | -!- FUNCTION: SEED STORAGE PROTEIN. | | | | |
| 12 69 19.4 239 1 WTL_SMMIMA | P09953 | smintthopsis | CC | CC | BODIES | | | | |
| 13 69 19.4 288 1 WTL_ALIMI | P09020 | alligator m | CC | CC | -!- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN, CONVICILIN, CONGYCINTIN, ETC.). | | | | |
| 14 69 19.4 448 1 WTL_RAT | P19952 | rattus norv | CC | CC | CONVICILIN, CONGYCINTIN, ETC.). | | | | |
| 15 69 19.4 449 1 WTL_HUMAN | P19544 | homo sapien | CC | CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | | | | |
| 16 69 19.4 449 1 WTL_MOUSE | P22561 | mus musculu | CC | CC | -!- SUBCELLULAR LOCATION: COTYLEDONARY MEMBRANE-BOUND VACUOLAR PROTEIN | | | | |
| 17 69 19.4 449 1 WTL_PIG | Q02651 | sus scrofa | CC | CC | BODIES | | | | |
| 18 69 19.4 895 1 ANDR_PAPHA | Q97960 | papio hamad | CC | CC | -!- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN, CONVICILIN, CONGYCINTIN, ETC.). | | | | |
| 19 69 19.4 1023 1 GLT_DROME | P33438 | drosophila | CC | CC | CONVICILIN, CONGYCINTIN, ETC.). | | | | |
| 20 68.5 19.3 758 1 YM38_YEAST | P03885 | saccharomye | CC | CC | CONVICILIN, CONGYCINTIN, ETC.). | | | | |
| 21 68.5 19.3 1382 1 IF3A_HUMAN | P14152 | homo sapien | CC | CC | CONVICILIN, CONGYCINTIN, ETC.). | | | | |
| 22 68 19.2 1154 1 WCL_NEUCR | P01371 | neurospora | FT | FT | INTERPRO: IPR00113; PF00546; Seedsstore-7s; 1. | | | | |
| 23 68 19.2 1265 1 CY45_CANFA | P0803 | cannis famili | FT | FT | KW | | | | |
| 24 67.5 19.0 164 1 2553_ARATH | P15459 | arabidopsis | FT | FT | Seed storage protein; signal. | | | | |
| 25 66.5 18.7 170 1 2522_ARATH | P15458 | arabidopsis | FT | FT | SIGNAL | | | | |
| 26 66.5 18.7 372 1 HX12_MOUSE | P11245 | mus musculu | FT | FT | CHAIN | | | | |
| 27 66.5 18.7 372 1 HX12_RAT | P31246 | rattus norv | FT | FT | SEQUENCE | | | | |
| 28 66 18.6 423 1 TFR_HUMAN | Q22664 | homo sapien | SQ | SQ | 588 AA; 69729 MW; 63E699B29AB8ADEB CRC64; | | | | |
| 29 66 18.6 529 1 GAG_MLVDE | P23090 | duplicl muri | | | Query Match | | | | |
| 30 66 18.6 536 1 GAG_MLVDE | P29168 | murine leuk | | | Best Local Similarity | | | | |
| 31 66 18.6 538 1 GAG_MLVFP | P26805 | friend muri | | | 33; Conservative | | | | |
| 32 66 18.6 539 1 GAG_MLVFP | P26807 | friend muri | | | Mismatches | | | | |
| 33 66 18.6 540 1 GAG_MLVHO | P21435 | homyl muri | | | 18; Indels | | | | |
| | | | | | 38; Gaps | | | | |
| | | | | | 5; | | | | |
| | | | | | QY | 3 DPQOREYEDRRCEEQEQPQQHQQLRC-----REQQ-----ROHGRGGDM-----45. | | | |
| | | | | | Db | 82 DPQRR-YEECQECRQERQRQPQCQCRKRFQEQQQSQSRQFOECQQHQHQEQPER 140 | | | |
| | | | | | OY | 46 -----NPQR-----GGSGRYEEGERES 63 | | | |
| | | | | | Db | 141 KQQCVERCREERYQENPWRREEEAEETEEEGEQQS 178 | | | |

RESULT 2
 VCLA-GOSHI STANDARD; PRT; 605 AA.
 ID VCLA-GOSHI STANDARD; PRT; 605 AA.
 AC P09759;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE VICILIN GC72-A PRECURSOR (ALPHA-GLOBULIN A).
 OS Gossypium hirsutum (Upland cotton).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicots; Rosidae; euroids II;
 OC Malvales; Malvaceae; Gossypium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chilan C.A., Borroto K., Kamalay J.A., Dure L. III;
 RT "Developmental biochemistry of cottonseed embryogenesis and
 germination. XIX. Sequences and genomic organization of the alpha
 globulin (vicilin) genes of cottonseed.",
 RL Plant Mol. Biol. 9:533-546(1987).
 CC -!
 CC -! SUBCELLULAR LOCATION: COTYLEDONARY MEMBRANE-BOUND VACUOLAR PROTEIN
 CC -!
 CC -! SIMILARITY: TO OTHER TS SEED STORAGE PROTEINS (PHASOLIN, VICILIN,
 CC CONVICTILIN, CONGLYCININ, ETC.).
 CC -!
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial/
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -!
 CC DR PIR: M19378; AAA33059.1; -.
 DR HSSP: P50477; ICAX.
 DR INTERPRO: IPR001113; -.
 DR PF00546; Seedstore_7s; 1.
 KW Seed storage protein; Signal.
 FT SIGNAL 1
 FT CHAIN 24 605 VICILIN GC72-A.
 SQ SEQUENCE 605 AA; 71049 MW; C9DB9371C9/6953B CRC64;

Query Match 28.5%; Score 101; DB 1; Length 605;
 Best Local Similarity 27.7%; Pdb. No. 0; 0.002;
 Matches 26; Conservative 15; Mismatches 19; Indels 34; Gaps 4;
 Oy 3 DPOQREYEDCRRCCQEPROOHOQLRCQ-OHQHGRGGMM..... 45
 |||: 1:1:1: 1:1:1 | : 1:1:1 :1 : 1:
 Db 79 DPQR-YDCCRQHQEERRLRPHQCSREQYERQQQQPDKQRECOORCQWQBQRP 137
 Oy 46 -! NPQRG-GSGRVEEGEEQS 63
 : 1 :| : 1 :1:1 : 1 1:1:1
 Db 138 RKQCVKECREBOYEDPWKGGERENKWRREEEEE 171

RESULT 3
 SSN6 YEAST STANDARD; PRT; 966 AA.
 ID SSN6 YEAST STANDARD; PRT; 966 AA.
 AC P14922;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE GLUCOSE REPRESSION MEDIATOR PROTEIN.
 GN SSN6 OR CYC8 OR YBL12C OR YBR0908.
 OS Saccharomycetaceae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomyctetes; Saccharomycetales;
 OC Saccharomyctaceae; Saccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A.
 MEDLINE=89211964; PubMed=2854095;

RA Trumbly R.J.;
 RT "Cloning and characterization of the CYC8 gene mediating glucose
 repression in yeast.";
 RL Gene 73:97-111(1988).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88065502; PubMed=3316983;
 RA Schulz J., Carlson M.;
 RT "Molecular analysis of SSN6, a gene functionally related to the SNP1
 protein kinase of *Saccharomyces cerevisiae*.";
 RL Mol. Cell. Biol. 7:3637-3645(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=92327848; PubMed=1626431;
 RA Mannhaupt G., Stucka R., Bhlle S., Vetter I., Feldmann H.;
 RT "Molecular analysis of yeast chromosome II between CMD1 and LYS2: the
 excision repair gene RAD6 located in this region belongs to a novel
 group of double-finger proteins.";
 RL Yeast 8:397-408(1992).
 RN [4]
 RP REPEATS.
 RX MEDLINE=90124639; PubMed=2404612;
 RA Sikorski R.S., Boguski M.S., Goebel P.A.;
 RT "A repeating amino acid motif in Cdc23 defines a family of proteins
 and a new relationship among genes required for mitosis and RNA
 synthesis.";
 RL Cell 60:307-317(1990).
 CC -!
 CC -! FUNCTION: IT IS INVOLVED IN REPRESSION BY A1-ALPHA2 AND ALPHA2 AND
 IN OTHER SYSTEMS AS A GENERAL REPRESSOR OF TRANSCRIPTION. THIS
 PROTEIN HAS NO OBVIOUS DNA-BINDING DOMAINS. IT MIGHT NOT INTERACT
 DIRECTLY WITH DNA BUT WITH DNA-BOUND PROTEINS.
 CC -!
 CC -! SUBCELLULAR LOCATION: NUCLEAR.
 CC -!
 CC -! SIMILARITY: CONTAINS 10 TPR DOMAINS.
 CC -!
 CC -! SIMILARITY: TO YEAST Cdc11 AND CCR4.
 CC -!
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial/
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -!
 CC DR EMBL; M23440; AAA34545.1; -.
 DR EMBL; M17826; AA35103.1; -.
 DR EMBL; X66247; CAA46973.1; -.
 DR EMBL; X78993; CAK5615.1; -.
 DR EMBL; Z35981; CAK5009.1; -.
 DR PIR; S25365; S25365.
 DR SGD; S000316; CYC8.
 DR INTERPRO: IPR001440; -.
 DR PFAM: PF00515; TPR; 7.
 DR KW Transcription regulation; Repressor; Repeat; TPR domain;
 KW Nuclear protein.
 FT DOMAIN 15 30 POLY-GUN.
 FT REPEAT 46 79 TPR 1.
 FT REPEAT 80 113 TPR 2.
 FT REPEAT 114 147 TPR 3.
 FT REPEAT 150 183 TPR 4.
 FT REPEAT 187 220 TPR 5.
 FT REPEAT 224 257 TPR 6.
 FT REPEAT 258 291 TPR 7.
 FT REPEAT 296 329 TPR 8.
 FT REPEAT 330 363 TPR 9.
 FT REPEAT 364 398 TPR 10.
 FT DOMAIN 493 556 30 X 2 AA TANDEM REPEATS OF Q-A.
 FT DOMAIN 557 587 POLY-GLN.
 FT CONFLICT 547 547 K -> Q (IN REF. 3).
 SQ SEQUENCE 966 AA; 107202 MW; 84B509CF320BC5C0 CRC64;

Query Match 23.7%; Score 84; DB 1; Length 966;

[17] VARIANT CATS ASN-695 AND HIS-695, AND SEQUENCE OF 629-723 FROM N.A.
 RP RESULT 5
 RX NTIT4_NEUCR
 RA ID NTIT4_NEUCR
 RA STANDARD; PRT; 1090 AA.
 RA P28349;
 RA "Ris-Stalpers C., Trifiro M.A., Kuiper G.G., Jenster G., Romalo G.,
 RT P28349;
 RT "Substitution of aspartic acid 686 by histidine or asparagine in the
 RT human androgen receptor leads to a functionally inactive protein with
 RT altered hormone-binding characteristics.";
 RL Mol. Endocrinol. 5:1562-1569(1991).
 RN [18]
 RP VARIANT CATS AND PAIS.
 RX MEDLINE=93338440; PubMed=1307250;
 RA batch J.A., Williams D.M., Davies H.R., Brown B.D., Evans B.A.J.,
 RA Hughes I.A., Patterson M.N.;
 RT "Androgen receptor gene mutations identified by SSCP in fourteen
 RT subjects with androgen insensitivity syndrome.";
 RL Hum. Genet. 1:497-503(1992).
 RN [19]
 RP VARIANT CATS VAL-787.
 RX MEDLINE=92235226; PubMed=1569163;
 RA Nakao R., Haji M., Yanase T., Ogo A., Takayanagi R., Katsube T.,
 RA Fukumaki Y., Navata H.;
 RT "A single amino acid substitution (Met-86-->Val) in the steroid-
 RT binding domain of human androgen receptor leads to complete androgen
 RT insensitivity syndrome.";
 RL J. Clin. Endocrinol. Metab. 74:1152-1157(1992).
 RN [20]
 RP VARIANT LNCAP ALA-877.
 RX MEDLINE=9222955; PubMed=1562539;
 RA Veldscholte J., Berrevoets C.A., Ris-Stalpers C., Kuiper G.G.,
 Jenster G., Trapman J., Brinkmann A.O., Mulder E.;
 RT "The androgen receptor in LNCAP cells contains a mutation in the
 ligand binding domain which affects steroid binding characteristics
 and response to antiandrogens.";
 RL J. Steroid Biochem. Mol. Biol. 41:665-669(1992).
 RN [21]
 RP VARIANT MET-730.
 RX MEDLINE=92135289; PubMed=1631125;
 RA Newmark J.R., Hardy D.O., Tomb D.C., Carter B.S., Epstein J.L.,
 Isaacs W.B., Brown T.R., Barrack E.R.;
 RT "An exonic point mutation creates a MaeII site in the androgen
 RT receptor gene mutations in human prostate cancer.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6319-6323(1992).
 RN [22]
 RP VARIANT CATS VAL-754.
 RX MEDLINE=93372805; PubMed=8103938;
 RA Lobaccaro J.-M., Lumbruso S., Ktari R., Dumas R., Sultan C.;
 RA "An exonic point mutation creates a MaeII site in the androgen
 RT receptor gene of a family with complete androgen insensitivity
 RT syndrome.";
 RL Hum. Mol. Genet. 2:1041-1043(1993).
 RN [23]
 RP VARIANT CATS ARG-807.
 RX MEDLINE=94108430; PubMed=8281140;
 RA Adeyemo O., Kallio P.J., Palvinen J.J., Kontula K., Jaenne O.A.;
 RT "A single base substitution in exon 6 of the androgen receptor gene
 RT causing complete androgen insensitivity: the mutated receptor fails
 RT to transactivate but binds to DNA in vitro.";
 RL Hum. Mol. Genet. 2:1809-1812(1993).
 RN [24]
 RP VARIANT PAIS VAL-743.
 RX MEDLINE=9331566; PubMed=8325932;
 RA Nakao R., Yanase T., Sakai Y., Haji M., Nawata H.;
 RT "A single amino acid substitution (Gly743 --> Val) in the steroid-
 binding domain of the human androgen receptor leads to Reiffenstein

Query Match 22.3%; Score 79; DB 1; Length 919;

Best Local Similarity 32.8%; Pred. No. 0.51; Gaps 1;

Matches 19; Conservative 15; Mismatches 22; Indels 2;

QY 5 QOREVEDCRRCEQQEPHQHQCQLRCREOORHQGRGGDMNPQQGGSGRYEEGEGEQ 62
 Db 58 QQQQQQQ-QQQQQQQQQQQQQQQQQSPRQQQQQQQQEDGSQAHRGCGPTGVLVLEBQQ 113

RESULT 5
 RP NTIT4_NEUCR
 RX ID NTIT4_NEUCR
 RA STANDARD; PRT; 1090 AA.
 RA P28349;
 RA "01-DEC-1992 (Rel. 24, Created)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE NITROGEN ASSIMILATION TRANSCRIPTION FACTOR NTIT-4.
 GN NTIT-4.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae; Neurospora.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92017855; PubMed=1840634;
 RA Yuan G.-F., Fu Y.-H., Marzluf G.A.;
 RT "nit-4, a pathway-specific regulatory gene of Neurospora crassa,
 encodes a protein with a putative binuclear zinc DNA-binding
 domain.";
 RL Mol. Cell. Biol. 11:5735-5745(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92149315; PubMed=1531376;
 RA Yuan G.-F., Marzluf G.A.;
 RT "Molecular characterization of mutations of nit-4, the
 nitrate-specific regulatory gene which controls nitrate assimilation
 in Neurospora crassa.";
 RL Mol. Microbiol. 6:67-73(1992).
 CC -I- FUNCTION: PATHWAY-SPECIFIC REGULATORY GENE OF NITRATE
 ASSIMILATION; IT ACTIVATES THE TRANSCRIPTION OF THE GENES FOR
 CC NITRATE AND NITRATE REDUCTASES.
 CC -I- SUBCELLULAR LOCATION: NUCLEAR.
 CC -I- DOMAIN: THE GLUTAMINE-RICH DOMAIN MIGHT FUNCTION IN ACTIVATING
 CC GENE EXPRESSION.
 CC -I- SIMILARITY: CONTAINS A ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
 CC CLUSTER DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M80368; AAA33602.1; -.
 DR PIR; A41696; A41696.
 DR PIR; S20033; S20033.
 DR HSSP; P07272; IPMI.
 DR INTERPRO; IPR001138; -.
 DR PFAM; PF00172; ZN_CluS; 1.
 DR PROSITE; PS50463; ZN2_CYS6_FUNGAL_1; 1.
 DR KW Transcription regulation; Activator; DNA-binding; Nuclear protein;
 KW Zinc; Metal-binding; Nitrate assimilation. ZN(2)-CYS(6), FUNGAL-TYPE.
 FT DNA_BIND 53 81 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 121 139 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 213 229 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 429 450 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 672 754 PRO-RICH.
 FT DOMAIN 755 859 GLN-RICH.
 FT DOMAIN 992 1024 POLY-GLN.
 FT CONFLICT 98 99 K->KP (IN REF. 1).
 FT CONFLICT 467 467 L->S (IN REF. 1).
 SQ SEQUENCE 1090 AA; 12024 MW; 881D89172EDD6114 CRC64;

Query Match 21.1%; Score 75; DB 1; Length 1090;

Best Local Similarity 32.7%; Pred. No. 1.5; Gaps 1;

Matches 17; Conservative 14; Mismatches 17; Indels 4; Gaps 1;

QY 6 QREYEDCRRCEQQEPHQHQCQLRCREOORHQGRGGDMNPQQGGSGRYEEGEGEQ 53
 Db 58 QQQQQQQ-QQQQQQQQQQQQQQQQQSPRQQQQQQQQEDGSQAHRGCGPTGVLVLEBQQ 113

OS Alligator mississippiensis (American alligator).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
 RN [1]
 RP SEQUENCE FROM N A.
 RX MEDLINE=96068905; PubMed=7478606;
 RT Kent J., Corrit A.M., Sharpe P.T., Hastie N., van Heyningen V.;
 "The evolution of Wt1 sequence and expression pattern in the
 vertebrates";
 Oncogen 11:1781-1792(1995).
 RL -I- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
 AND BINDS TO THE DNA SEQUENCE 5'-CGCCCGCC-3'.
 CC !- SUBCELLULAR LOCATION: NUCLEAR.
 CC !- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS (BY
 SIMILARITY).
 CC !- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC FINGER
 PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 of the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcements>
 or send an email to license@isb-sib.ch).
 DR
 DR EMBL: X85730; CAA59735.1; -.
 DR HSSP: P08046; IALL.
 DR INTERPRO: IPR000822; -.
 DR PFAM: PF00096; zf-C2H2; 4.
 DR PROSTRE: PS0028; ZINC_FINGER_C2H2; 4.
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
 KW Transcription regulation; Alternative splicing.
 FT NON_TER 1 1
 FT ZN_FING 162 186 C2H2-TYPE.
 FT ZN_FING 192 216 C2H2-TYPE.
 FT ZN_FING 222 244 C2H2-TYPE.
 FT ZN_FING 253 277 C2H2-TYPE.
 FT SEQUENCE 245 247 MISSING (IN ISOFORM 2).
 SQ 288 AA; 3311 MW; 33E2F7DBE7BFDE CRC64;
 RESULT 14
 WTL-RAT
 ID WTL-RAT STANDARD; PRT; 448 AA.
 AC P49552;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE WILMS' TUMOR PROTEIN HOMOLOG
 WT1 OR WT-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N A.
 STRAIN=DAWLEY; TISSUE=KIDNEY;
 RX MEDLINE=93046155; PubMed=1330293;
 RA Sharma P.M., Yang X., Bowman M., Roberts V., Sukumar S.;
 "Molecular cloning of rat Wilms' tumor complementary DNA and a study
 of messenger RNA expression in the urogenital system and the brain.",
 Cancer Res. 52:6407-6412(1992).
 CC !- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
 AND BINDS TO THE DNA SEQUENCE 5'-CGCCCGCC-3'.

| | | |
|---|---|--|
| CC | CC | -.- SUBCELLULAR LOCATION: NUCLEAR. |
| CC | CC | -.- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS. |
| CC | CC | -.- TISSUE SPECIFICITY: KIDNEY (BY SIMILARITY). |
| CC | CC | -.- DEVELOPMENTAL STAGE: EXPRESSED DURING KIDNEY DEVELOPMENT. |
| CC | CC | -.- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS. |
| CC | CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). |
| CC | CC | EMBL: X69716; CAA49373.1; -. |
| DR | DR | TRANSFAC: T0252; |
| DR | DR | INTERPRO: IPR00822; -. |
| DR | DR | INTERPRO: IPR00976; -. |
| DR | DR | Pfam: PF0096; zf-C2H2; 4. |
| DR | DR | PRINTS: PR00048; ZINC-FINGER. |
| DR | DR | PROSITE: PS00028; ZINC-FINGER.C2H2; 4. |
| KW | KW | Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein; Transcription factor; Zinc-finger; Zinc-finger; Zinc-finger; Zinc-finger. |
| FT | FT | DOMAIN 27 82 PRO-RICH. |
| FT | FT | DOMAIN 322 437 ZINC FINGERS. |
| FT | FT | ZN_FING 322 346 C2H2-TYPE. |
| FT | FT | ZN_FING 352 376 C2H2-TYPE. |
| FT | FT | ZN_FING 382 404 C2H2-TYPE. |
| FT | FT | ZN_FING 413 437 C2H2-TYPE. |
| FT | FT | VARSPLIC 249 265 MISSING (IN ISOFORM 2 AND ISOFORM 3). |
| FT | FT | VARSPLIC 407 409 MISSING (IN ISOFORM 2 AND ISOFORM 4). |
| SQ | SQ | 448 AA; 49193 MW; 324AC9AC1FF73F76 CRC64; |
| Query Match | | |
| Best Local Similarity 19 4%; Score 69; DB 1; Length 448; | | |
| Matches 15; Conservative 14; Mismatches 25; Indels 6; Gaps 1; | | |
| QY | ID | POOREYEDCCRRCQEQPQQHQOC-----LRCEBOORQHGRRGGMMNPORGSGRYEE 57 |
| Db | AC | P19544; 01-FEB-1991 (Rel. 17, created) |
| | AC | 01-AUG-1991 (Rel. 19, last sequence update) |
| | DT | 01-OCT-2000 (Rel. 40, last annotation update) |
| | DE | WILMS' TUMOR PROTEIN (WT133). |
| GN | WT1. | |
| OS | Homo sapiens (Human). | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| RN | [1] | |
| RP | SEQUENCE FROM N.A. | |
| RC | TISSUE-PETAL KIDNEY; | |
| RC | MEDLINE:90158822; PubMed=2154702; | |
| RX | Gossler M., Pousta A., Cavenee W., Neve R. L., Orkin S. H., | |
| RA | Bruns G. A. P.; Haber D. A., Sohn R. L., Buckler A. J., Pelletier J., Call K. M., | |
| RA | Housman D. E.; Alternative splicing and genomic structure of the Wilms tumor gene | |
| RT | "Homozygous deletion in Wilms tumours of a zinc-finger gene identified by chromosome jumping."; | |
| RT | Nature 343:774-778(1990). | |
| RL | SEQUENCE FROM N.A. | |
| RN | TISSUE-PLACENTA; | |
| RX | MEDLINE:92052142; PubMed=1658787; | |
| RA | Haber D. A., Sohn R. L., Buckler A. J., Pelletier J., Call K. M., | |
| RT | "Alternative splicing and genomic structure of the Wilms tumor gene | |

RT Wt1.";
 RL proc. Natl. Acad. Sci. U.S.A. 88:9618-9622(1991).
 RN [3]
 RP SEQUENCE OF 85-449 FROM N A.
 RX MEDLINE=90150277; PubMed=2154335;
 RA Call K.M., Glaser T., Ito C.Y., Buckler A.J., Pelletier J.,
 RA Haber D.A., Rose E.A., Kral A., Yeger H., Lewis W.H., Jones C.,
 RA Housman D.E.,
 RT "Isolation and characterization of a zinc finger polypeptide gene at
 the human chromosome 11 Wilms' tumor locus."
 RL Cell 60:509-520(1990).
 RN
 RP IDENTIFICATION OF START CODON AND ALTERNATIVE SPLICE SITES.
 RX MEDLINE=9141522; PubMed=1671709;
 RA BUCKLER A.J., Glaser T., Haber D.A., Glaser T., Housman D.E.;
 RT "Isolation, characterization, and expression of the murine Wilms'
 RT tumor gene (Wt1) during kidney development."
 Mol. Cell. Biol. 11:1707-1712(1991).
 RL
 RN [5]
 RP VARIANT WT CYS-366.
 RX MEDLINE=92279213; PubMed=1317572;
 RA Little M.H., Prosser J., Condie A., Smith P.J., van Heyningen V.,
 RA Hastic N.D.;
 RT "Zinc finger point mutations within the Wt1 gene in Wilms tumor
 patients."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:4791-4795(1992).
 RN
 RP VARIANT DDS.
 RX MEDLINE=92005721; PubMed=1655284;
 RA Peilletter J., Bruehnen W., Kashian C.E., Mauer S.M., Manivel J.C.,
 RA Striebel J.E., Houghton D.C., Junien C., Habib R., Fouser L.,
 RA Finn R.N., Silverman B.L., Haber D.A., Housman D.;
 RT "Germline mutations in the Wilms' tumor suppressor gene are
 RT associated with abnormal urogenital development in Denys-Drash
 syndrome."
 RL Cell 67:437-447(1991).
 RN [7]
 RP VARIANT DDS.
 RX MEDLINE=93265053; PubMed=1338906;
 RA Baird P.N., Santos A., Groves N., Jadresic L., Cowell J.K.;
 RT "Constitutional mutations in the Wt1 gene in patients with
 RT Denys-Drash syndrome."
 RL Hum. Mol. Genet. 1:301-305(1992).
 RN [8]
 RP VARIANT DDS.
 RX MEDLINE=93219183; PubMed=6388765;
 RA Little M.H., Williamson K.A., Mannens M., Kelsey A., Gosden C.,
 RA "Hastie N., van Heyningen V.,
 RT "Evidence that Wt1 mutations in Denys-Drash syndrome patients may act
 in a dominant-negative fashion."
 Hum. Mol. Genet. 2:2559-264(1993).
 RN [9]
 RP VARIANT MESOTHELIOMA GLY-273.
 RX MEDLINE=94004972; PubMed=8401592;
 RA Park S., Schalling M., Bernard A., Maheswaran S., Shipley G.C.,
 RA Roberts D., Fleicher J., Shipman R., Rheinwald J., Demetri G.,
 RA Griffin J., Minden M., Housman D.E., Haber D.A.;
 RT "The Wilms' tumour gene Wt1 is expressed in murine mesoderm-derived
 tissues and mutated in a human mesothelioma."
 Nat. Genet. 4:415-420(1993).
 RL
 RN [10]
 RP VARIANT WT SER-181 AND ALA-253.
 RX MEDLINE=97268881; PubMed=9108099;
 RA Schumacher V., Schneider S., Fligge A., Wildhardt G., Harms D.,
 RT "Correlation of germ-line mutations and two-hit inactivation of the
 Wt1 gene with Wilms' tumors of stromal-predominant histology."
 RL Proc. Natl. Acad. Sci. U.S.A. 94:3972-3977(1997).
 RN [11]
 RP VARIANTS DMS TYR-377; LEU-383 AND ASN-396.
 RX MEDLINE=98198341; PubMed=9529164;
 RA Jeampierre C., Denamur E., Henry I., Cabanis M.-O., Luce S.,
 Cecile A., Elion J., Peuchmaur M., Loirat C., Niaudet P.,

RA Gubler M.-C., Junien C.;
 RT "Identification of constitutional Wt1 mutations, in patients with
 isolated diffuse mesangial sclerosis, and analysis of
 genotype/phenotype correlations by use of a computerized mutation
 database."
 RL Am. J. Hum. Genet. 62:824-833(1998).
 RN [12]
 RP REVIEW.
 RX MEDLINE=92207913; PubMed=1313285;
 RA Haber D.A., Buckler A.J.;
 RT "Wt1: a novel tumor suppressor gene inactivated in Wilms' tumor."
 RL New Biol. 4:97-106(1992).
 RN [13]
 RP REVIEW.
 RX MEDLINE=93345769; PubMed=8393820;
 RA Rauscher F.J. III;
 RT "The Wt1 Wilms' tumor gene product: a developmentally regulated
 transcription factor in the kidney that functions as a tumor
 suppressor."
 RL PASEB J. 7:896-903(1993).
 CC -I- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
 AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCGC-3'.
 CC -I- SUBCELLULAR LOCATION: NUCLEAR.
 CC -I- ALTERNATIVE PRODUCTS: TWO ALTERNATIVE SPLICE SITES EXISTS.
 CC -I- TISSUE SPECIFICITY: EXPRESSED IN THE KIDNEY AND A SUBSET OF
 HEMATOPOIETIC CELLS.
 CC -I- DISEASE: WILM'S TUMOR (WT) IS AN EMBRYONAL MALIGNANCY OF THE
 KIDNEY THAT AFFECTS APPROXIMATELY 1 IN 10,000 INFANTS AND YOUNG
 CHILDREN. IT OCCURS BOTH IN SPORADIC AND HEREDITARY FORMS.
 CC -I- INACTIVATION OF Wt1 IS ONE OF THE CAUSES OF WILM'S TUMOR.
 CC -I- DISEASE: PATIENTS WITH DENYS-DRASH SYNDROME (DDS), WHICH IS
 CHARACTERIZED BY A TYPICAL NEPHROPATHY AND GENITAL ABNORMALITIES,
 HAVE DEFECTS IN THE Wt1 GENE.
 CC -I- DISEASE: DEFECTS IN Wt1 ARE ALSO A CAUSE OF DIFFUSE MESANGIAL
 SCLEROSIS (DMS), A FORM A DDS.
 CC -I- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
 PROTEINS.
 CC -I- DATABASE: NAME=Wt1; NOTE=Wt1 mutation database;
 CC WWW="http://www.undec.noekker.fr/203/".

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

 DR X53596; XAA61299.1; ALT-INIT.
 EMBL: M8032; XAA61299.1;
 DR EMBL: M80217; XAA61299.1; JOINED.
 DR EMBL: M80218; XAA61299.1; JOINED.
 DR EMBL: M80219; XAA61299.1; JOINED.
 DR EMBL: M80220; XAA61299.1; JOINED.
 DR EMBL: M80221; XAA61299.1; JOINED.
 DR EMBL: M80228; XAA61299.1; JOINED.
 DR EMBL: M80229; XAA61299.1; JOINED.
 DR EMBL: M80231; XAA61299.1; JOINED.
 DR EMBL: M30593; XAA61299.1; JOINED.
 DR PIR: A34673; A34673.
 DR PIR; S08273; S08273.
 DR TRANSAC; T00899; -.
 DR MIM: 194070; -.
 DR MIM: 194080; -.
 DR MTM: 255370; -.
 DR INTERPRO; IPR000822; -.
 DR INTERPRO; IPR000976; -.
 DR PFAM; PF00096; zf-C2H2; 4.
 DR PRINTS; PR00048; ZINCFINER.
 DR PRINTS; PR00049; WILMSTUMOUR.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
 KW Transcription regulation; Alternative splicing; Anti-oncogene;

KW
Disease mutation.
FT DOMAIN 27 83 PRO-RICH.
FT DOMAIN 323 438 ZINC FINGERS.
FT ZN_FING 323 347 C2H2-TYPE.
FT ZN_FING 353 377 C2H2-TYPE.
FT ZN_FING 383 405 C2H2-TYPE.
FT ZN_FING 414 438 C2H2-TYPE.
FT VARSPLIC 250 266 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPLIC 408 410 MISSING (IN ISOFORM 2 AND ISOFORM 4).
FT VARIANT 181 181 P -> S (IN WT).
FT VARIANT 223 223 /FTId=VAR_007739.
FT VARIANT 253 253 S -> N (IN WT).
FT VARIANT 273 273 G -> A (IN WT).
FT VARIANT 330 330 /FTId=VAR_007740.
FT VARIANT 360 360 S -> G (IN WT).
FT VARIANT 366 366 /FTId=VAR_007741.
FT VARIANT 366 366 S -> G (IN MESOTHELIOMA).
FT VARIANT 366 366 /FTId=VAR_007742.
FT VARIANT 366 366 C -> Y (IN DDS).
FT VARIANT 366 366 /FTId=VAR_007743.
FT VARIANT 373 373 C -> G (IN DDS).
FT VARIANT 373 373 /FTId=VAR_007744.
FT VARIANT 377 377 R -> C (IN WT).
FT VARIANT 377 377 /FTId=VAR_007745.
FT VARIANT 377 377 H -> R (IN DDS).
FT VARIANT 377 377 /FTId=VAR_007746.
FT VARIANT 377 377 O -> H (IN DDS).
FT VARIANT 377 377 /FTId=VAR_007747.
FT VARIANT 377 377 Y -> H (IN DDS/DMs).

Query Match: 19.4%; Score: 69; DB: 1; Length: 449;
Best Local Similarity: 25.0%; Pred. No.: 2.7;
Matches: 15; Conservative: 14; Mismatches: 25; Indels: 6; Gaps: 1;

QY 4 PPOREYEDCRRRQEQQEPROHQO-----LRCREGOROHGRGGDMNPQRGSGREE 57
Db 352 PYQCDFKOCERRRSRSDQLKRRHRTGKVPKTQCKTCQRKFSSRSDDHKTHTRHGTKE 411

Search completed: March 1, 2001, 16:16:33
Job time: 413 sec

